Optimal 3D Viewing with Adaptive Stereo Displays for Advanced Telemanipulation

Sukhan Lee[1,2] Srinivasan Lakshmanan[1] Sookwang Ro[1]
Jong-Oh Park[3] Chong-Won Lee[3]

Dept. of EE-Systems and CS[1]
University of Southern California
Los Angeles, CA 90089-0781

Jet propulsion Laboratory[2]
California Institute of Technology
Pasadena, CA 91109

Korea Institute of Science[3]
and Technology
Seoul, Korea, 136-791

Abstract

A method of optimal 3D viewing base.d on adaptive displays of stereo images is presented for advanced telemanipulation. The method provides the viewer with the capability of accurately observing a virtual 31) object or local scene of his/her choice with minimum distortion. The viewer is allowed to define a virtual 3D object or local scene at a desired depth as a scaled version of a real 31) object or local scene which he/she wants to focus on. The key result is an algorithm which determines in real-time the optimal parameter values associated with stereo imaging and image projection on a video screen. The selected parameter values implements the optimal stereo viewing adaptive to the viewer's focus of attention. Simulation results arc shown.

1. introduction

Telepresence is an essential component for telemanipulation[2]. Telepresence becomes more and more important as the advancement of telemanipulation demands highly dextrous and delicate manipulation of objects for such sophisticated tasks as telesurgery[1]. One of the key features in telepresence is to provide. the viewer with the capability of observing a 3D scene remotely as if he/she observe the 3D scene directly in the real world. in other words, to provide a human with the natural yet comfortable feeling of 3D visual perception based on stereo displays on a video screen(s) is important for advanced telemanipulation as well as virtual reality. More specifically, the image of a real 3D scene is taken by a stereo camera. And, the images recorded on the left and right camera image planes arc projected onto a video screen(s)[6] such that the viewer's left and right eyes scc the respective left and right camera images and synthesize the 3D scene by stereopsis[4]. In the case where the viewer wants to change his/her viewing perspective, the viewer's head orientation toward the desired viewing direction is captured by a he.acl-mounted device to control the orientation of stereo camera[3].

This paper concentrates on developing a method of stereo displays that provides the viewer with the capability of observing a 3D object or scene of his/her choice with high accuracy under the mag nification and depth set independently by the viewer. Previous work in this area includes shifting the left and right images in opposite dire.ctions to obtain undistorted stereo viewing with parallel cameras[5]. However, it requires that the system, not the viewer, controls the image size, which may not bc desirable for many applications, besides the suffering from a trade-off between depth resolution and stereo-image over-lap.

It is often the case that the viewer desire.s to set viewing parameters in such a way that he/she can observe a scaled version of the real 3D scene at a desired depth. Wc refer to bcrc tbc scaled version of a real 31) scene fictitiously placed at a desired depth from the viewer as a virtual 3D scene. Furthermore, the viewer may desire to focus on a particular object or a local scene at a time but with high accuracy without distortion. In fact, the behavior associated with the human recognition of an object is that the focus of human eyes fall on a particular feature of an object or scene at a time, as it jumps around among many features. The problem to solve here is how to set in real-time the parameters associated with stereo imaging as well as the parameters associated with the projection of stereo images on the screen such that the. viewer is capable of observing the object or local scene of his/her current focus of attention with desired magnification and depth feeling yet with minimum distortion, Cameras with wide and narrow fields of view with the ability to zoom arc an integral part of most camera systems. Here, wc assume that the controllable parameters associated with stereo imaging arc the baseline length and zoom factor or focal length in the case of pin-hole camera model. The controllable parameters associated with image projection on a screen are the horizontal displacement of stereo images in opposite direction and the image scaling factor. The image scaling factor and the camera zoom factor or focal length play an equivalent role. However, such redundancy helps to select parameters not only in terms of minimizing distortions but also satisfying other performance criteria or physical constraint associated with stereo viewing[5].

2. outline of the Approach

A virtual 3D scene defined by the viewer as a scaled version of a real 3D sccnc, which is placed at a desired depth from the viewer, may not be accurately synthesired by the eyes of the viewer without distortion by controlling only those imaging and projection parameters described above. Our approach is that we identify the object or local scene the viewer desires to focus on such that the parameter values arc determined adaptively in such a way that the eyes of the viewer can

synthesize the selected object or local scene optimally. For this purpose, we assume that the eye orientations can be effectively measured by a device such that the location of the viewer's current visual focus can be estimated in the virtual 3D visual world the viewer is currently engaged in. l'hen, by applying the desired magnification and depth scales set for the current virtual 3D visual world, the estimated location of the viewer's visual focus can be transformed into the corresponding location in the real 3D workspace. In some cases, it may be possible that the location of the object the viewer desires to track can be computed explicitly, for instance, the location of manipulator end-effector.

First, wc derive the equations that project a point in the real workspace onto the corresponding left and right points cm the screen through stereo imaging and image projection operations. These equations arc functions of controllable imaging and projection parameters. '1'hen, wc transform the point in the real workspace into the corresponding point in the virtual workspace based on the desire.d magnification and depth set by the viewer. That is, the virtual workspace is set as a scaled version of the real workspace located at a desired depth. Followed is the derivation of the equations that project the point in the virtual workspace onto the left and right points on the screen in such a way that the eyes of the viewer can perceive an object or scene with the desire.d magnification and depth and without distortion. The left and right points defined in the latter represent the ideal locations of a point. The optimization involves the minimization of error between the ideal locations of points from the virtual workspace. and the actual locations of points from the actual workspace by controlling the imaging and projection parameters. We propose two approaches for determining the optimal parameter values: First, based on the. estimated location of viewer's visual focus in the real workspace, wc select a number of points around the point of visual focus. Then, wc compute the optimal parameter values for each of the selected points individually, by taking advantage of the fact that there exists an exact solution for optimizing single point, i.e., wc can find the parameter values that allow a point in the virtual 3D scene exactly match with that of the corresponding real 3D scene. Then, final parameter values arc obtained by the weighted average of the individual parameter values, where the weights may be inversely proportional to the distance from the point of the viewer's visual focus. Second, wc. define the error functions as the squared sum of the errors from the individual points. An optimal solution which minimizes this error function is then derived to determine an optimal parameter values.

As an example, let us consider- that the viewer is given the 3D scene of manipulator workspace for telemanipulation. Assume that the viewer is curie.n(ly focusing on the manipulator end effector in the virtual workspace defined with the current magnification and depth set by the viewer. Now, the viewer wants to move his/her focus from the manipulator end-effector to the object to be grasped. The viewer then orients his/her eyes to the object in the virtual workspace, which is subsequentl y captured by the device measuring eye orientation. The viewer also controls a mouse to set the proper magnification and depth of the object in order to create a new virtual workspace. The location of the object in the real workspace is estimated from the estimated location of the object in the virtual workspace based on the measured eye orientations. q'hen, we select multiple points representing a volume around the estimated object location in the real workspace.. The final parameter values to be used for ncw stereo displays are then obtained from the weighted average of the parameter values computed for individual points or from the equations derived for multiple point optimization. The latter provides a more accurate solution at the expense of computation.

3. Projection of Real 3D World onto Screen

In this section, we derive the equations that project a point in the real 3D workspace. onto the left and right points on the video semen for the viewer's left and right eyes, respectively. 'l'his consists of two processes: 1) the stereo imaging in parallel cameras, and 2) the displays of stereo images on the screen.

3.1 Stereo Imaging.

As shown in Fig. 1, a point, P, in the real world, $P = (x, y, z)^T$ in the camera coordinate frame, forms the corresponding image points, P^{iL} and P^{iR}, 011 the left arid right camera image planes, respectively, where $P^{iL} = (P_x^{iL}, P_y^{iL})^T$ and $P^{iR} = (P_x^{iR}, P_y^{iR})^T$ in the respective left and right image plane frames. Assuming the parallel camera configuration with the pin-bole camera model, wc have

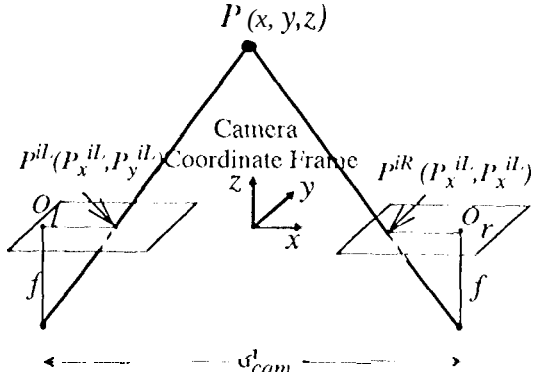

Fig 1. Stereo camera configuration

$$P_x^{iL} = \frac{f}{z+f}\left(x + \frac{d_{cam}}{2}\right) \quad (1)$$

$$P_x^{iR} = \frac{f}{z+f}\left(x - \frac{d_{cam}}{2}\right) \quad (2)$$

$$P_y^{iL} = P_y^{iR} = \frac{f}{z+f}y \quad (3)$$

where the baseline length of two parallel cameras, d_{cam}, and the zoom factor or the camera focal length, f,

are considered adjustable parameters.

3.2 Displaying Stereo Images on Screen

The left and right camera images arc them projected onto the video screen for stereo displays, where the projection involves the horizontal and vertical scaling of images with the scaling factors denoted respectively by S_x and S_y, and the horizontal shifting of images in the opposite direction with the amount of a symmetric shift denoted by K. The horizontal shift may be defined in terms of the screen coordinate frame located at the center of the screen. For convenience, here we ck.fine K in terms of the left and right screen frames which arc parallel to the screen frame and located directly in front of the left and right eyes of the viewer, respectively, as shown in Fig. 2. 'l'he eye frames of the viewer is assume.ct parallel to the screen frame and is located about the center of the screen. Then, the image points, P^{iL} and P^{iR}, arc projected onto the points, P^{SL} and P^{SR}, respectively, of the screen, where $P^{SL} = (Px^{SL}, P_y^{SL})^T$ and $P^{SR} = (Px^{SR}, P_y^{SR})^T$ can be represented in the respective left and right screen frames by

$$\begin{cases} P_x^{SL} = S_x \dfrac{f}{z+f}\left(x + \dfrac{1}{2}d_{cam}\right) - K \\ P_y^{SL} = S_y \dfrac{f}{z+f} y \end{cases} \quad (4)$$

$$\begin{cases} P_x^{SR} = S_x \dfrac{f}{z+f}\left(x - \dfrac{1}{2}d_{cam}\right) + K \\ P_y^{SR} = S_y \dfrac{f}{z+f} y \end{cases} \quad (5)$$

The scaling parameters, S_x, S_y, and the shifting parameters, K, arc considered adjustable. Eqs. (4) and (5) rc.present the projection of the point, P in the real world onto the points, P^{SL} and P^{SR} on the screen through stereo imaging and image projection, where P^{SL} and P^{SR} can be controlled by the adjustable parameters, d_{cam}, f, S_x, S_y, and K.

*iSL represents left projected image axis
*iSR represents right projected image axis

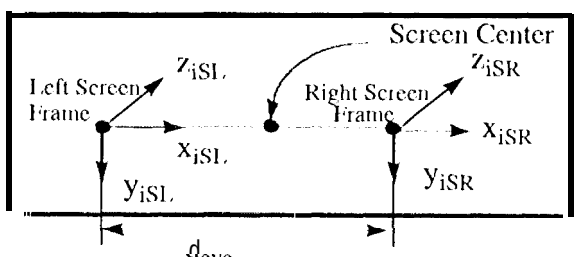

Fig. 2 Projection onto Screen

Note. that the adjustment of f is equivalent to adjusting zoom factor or field of view in the pin-hole camera model. Therefore, instead of performing digital image. scaling with S_x and S_y, wc can adjust f while keeping $S_x = S_y$ = screen size/image size. This can prevent the reduction of resolution duc to image magnification. However, to have S_x and S_y can help further minimize the viewing distortions.

4. Projection of Virtual 3D World onto Screen

A virtual 3D world is a scaled version of the corresponding real 3D world, which is placed at a desired depth from the viewer. A virtual world is defined by the viewer in order to observe an object of his/her interest with a desired magnification at a desired depth.

4.1 Generation of Virtual 3D Object

First, we imagine that the cyc frame of the viewer coincides with tbc camera frame, such that the real world seen by the camera is placed fictitiously in front of the viewer's eyes. 'l'hat is, a point, P, in the camera frame is represented as the corresponding point, P^e, $P^e = (x_e, y_e, z_e)$ in the eye frame, where $P = P^e$, as shown in Fig. 3. '[']hen, the Viewer wants to scale the real world and

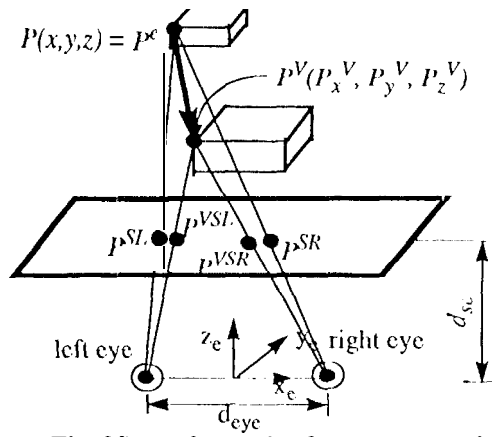

Fig. 3 Transformation from camera view to virtual view

place it at a certain depth from him/her. 'l'his can be done by scaling each point, P^e, with the desired scaling factor α and shifting the scaled point by t_z 10 the desired depth, such that P^e moves to the ncw position, P^V $P^V = (P_x^V, P_y^V, P_z^V)^T$:

$$P^V = \alpha(x,y,z)^T + (0,0,t_z)^T \quad (6)$$

4.2 Projection of' Virtual 3D Object onto Screen

Now, the virtual 3D world created by the viewer can be projected onto the screen so that the ideal position of P in the real 3D world can be defined on the screen. Similar to the case of stereo imaging, P^V can be projected onto the left and right points, P^{VSL} and P^{VSR}, respectively, on the screen. P^{VSL} and P^{VSR}, where P^{VSL} =

$(P_x^{VSL}, P_y^{VSL})^T$ and $P^{VSR} = (P_x^{VSR}, P_y^{VSR})^T$, can be defined in terms of the left and right screen frames, respectively, as follows:

$$P_x^{VSL} = \frac{\left(\alpha \cdot x + \frac{d_{eye}}{2}\right) \cdot d_{sc}}{\alpha \cdot z + t_z} \quad (7)$$

$$P_x^{VSR} = \frac{\left(\alpha x - \frac{d_{eye}}{2}\right) \cdot d_{sc}}{\alpha \cdot z + t_z} \quad (8)$$

and,

$$P_y^{VSL} = P_y^{VSR} = \frac{\alpha \cdot y \cdot d_{sc}}{\alpha \cdot z + t_z} \quad (9)$$

5. Selection of Optimal Parameters

The distortion in viewing is cause.d by the crmr between the screen points, P^{SL} and P^{SR}, projected from the stereo images and the screen points, P^{VSL} and P^{VSR}, projected from the virtual 3D world. To be free of any distortions, this error should be zero for each and every point cm the screen. Unfortunately, this is not achievable. only by adjusting the parameters, d_{cam}, f, S_x, S_y, and K for an arbitrary set of α and t_z.

Therefore, wc minimize the errors of selected points representing the volume where the viewer wants **to** focus cm, adaptively to the change of the viewer's visual focus of attention.

5.1 Distortions in Viewing

The eri ors between P^{SL} and P^{VSL} and between P^{SR} and P^{VSR}, repi esented as $(e_{xl}, e_{yl})^T$ and $(e_{xr}, e_{yr})^T$, respectively, can be obtained from Eqs, (4 - 9):

$$\begin{pmatrix} e_{xl} \\ e_{yl} \end{pmatrix} = \begin{pmatrix} P_x^{VSL} - P_x^{SL} \\ P_y^{VSL} - P_y^{SL} \end{pmatrix} \quad \textbf{(lo)}$$

$$\begin{pmatrix} e_{xr} \\ e_{yr} \end{pmatrix} = \begin{pmatrix} P_x^{VSR} - P_x^{SR} \\ P_y^{VSR} - P_y^{SR} \end{pmatrix} \quad (11)$$

5.2 Selection of Points to be Optimized

The selection of optimal parameters for the. minimization of viewing distortions is based on a number of selected points representing the volume around the viewer's cul-lent visual focus. Since human eyes can not sec. whole scene at a time, the scene dots not have to be optimized for all points. If optimization is done for whole scene then it causes more eri ors around the view point of human eye and it is not desirable. However enough points should be considered to get a sufficiently clear virtual scene.

The viewer's current visual focus, P^{vf}, in the virtual 3D world can be identified from the. measured orientations of viewer's eyes based on triangulation. 'I'hat is, by detecting the posture, orientation and distance between pupils of the viewer using a special device, his/her view point can be estimated. Then, P^{vf} can be transformed into the corresponding point, P^f, in the real 3D world by Eq. (6). A set of points is then selected from the volume. around P^f for optimization. The change of the viewer's visual focus prompts the selection of a new set of points for optimization.

5.3 Parameter Optimization

There arc two approaches for achieving parameter optimization with the selected points: 1) wc compute first the optimal parameter values for individual points. The final optimal parameter values are the weighted sum of individual optimal parameter values. This method is referred to here as the point-wise. optimization. 2) we compute the optimal parameter values to minimize the weighted sum of the squared errors from all the points. This method is referred to here as the multi-point optimization.

I'oirlt-Wise Optimization

Based on eqs. (10) and (11), the error, $(e_{xl}, e_{yl})^T$ and $(e_{xr}, e_{yr})^T$, representing the point-wise distortion of viewing can be made null for a given point by adjusting parameters based on:

$$S_x = S_y = \left(\frac{d_{sc}\alpha}{f}\right)\left(\frac{z+f}{\alpha z + t_z}\right) \quad \text{or}$$

$$f = \frac{d_{sc}\alpha z}{\alpha(z - d_{sc}) + t_z} \quad (12)$$

$$\left(\frac{\alpha d_{sc}}{\alpha z + t_z}\right)\left(d_{cam} - \frac{d_{eye}}{\alpha}\right) - 2K = 0 \quad (13)$$

The image scale factor, S_x and S_y, should be chosen same, Eq. (12) shows that the same result can be obtained by adjusting either the image scale factor, S_x or the focal length, f. Furthermore, Eq. (1 3) indicates that the selection of the camera baseline length. d_{cam} and the horizontal image shifting, K, **is** not unique. However, this redundancy in selecting parameter values arc rather desirable, since wc can taking into consideration additional performance criteria and physical constraints associated with S_x and f as well as d_{cam} and K. For instance, wc have a freedom to choose parameter values in such a way as to minimize the image resolution degradation, depth distortion, and stereo image overlap reduction. Note that the parameter values determined by Eqs. (1 2) and (13) arc only a function of the depth z, and arc independent of x and y.

Once the parameter values arc obtained for selected individual points, wc can obtain the fina I parameter values by the weighted average of individual values. The weights can be assigned to individual

points, P_i^f, inversely proportional to the distance from the reference point, P^f, as

$$\mathbf{W}_i \propto 1 / \| P_i^f - P^f \| \tag{14}$$

The result can be summarized as follows:

$$S_x^{opt} = \frac{\sum_i w_i S_x^i}{\sum_i w_i} \quad \text{or} \quad \frac{\sum_i w_i f^i}{\sum_i w_i} \tag{15}$$

$$K^{opt} = \frac{\sum_i w_i K^i}{\sum_i w_i} \tag{16}$$

$$d_{cam}^{opt} = \frac{\sum_i w_i d_{cam}^i}{\sum_i w_i} \tag{17}$$

The l~oir]t-wise'o~>till~ ization is simple to calculate and is suitable for real time computation.

Multi-Point Optimization

We first represent the total err or, ε, associated with the selected points as the weighted sum of squared errors of' individual points. From Eqs. (4- 11), ε can be represented as

$$\& = \varepsilon_y + \varepsilon_{xl} + \varepsilon_{xr} \tag{18}$$

$$\varepsilon_y = \sum_{i=1}^{N} \omega_i y_i (S_y b_i - a_i \alpha)^2$$

$$\varepsilon_{xl} = \sum_{i=1}^{N} \omega_i (x_i (S_x b_i - a_i \alpha) + c_i)^2$$

$$\varepsilon_{xr} = \sum_{i=1}^{N} \omega_i (x_i (S_x b_i - a_i \alpha) - c_i)^2$$

where

$$\begin{cases} a_i = \dfrac{d_{sc}}{\alpha z_i + t_z} \\ b_i = \dfrac{f}{z_i + f} \\ c_i = \dfrac{S_x b_i d_{cam}}{2} - \dfrac{a_i d_{eye}}{2} - K \end{cases}$$

From the necessary condition of optimality, i.e., the partial derivatives associated with parameters to be optimized should be zero, wc can obtain the following equations governing the optimal parameter values:

$$S_x (A + B d_{cam}^2) - d_{cam} (C + DK) = 0 \tag{19}$$

$$S_x d_{cam} \frac{D}{2} - KE = F \tag{20}$$

$$S_x d_{cam} \frac{B}{2} - K \frac{D}{2} = \frac{C}{2} \tag{21}$$

$$S_y = \frac{\sum_i 2\omega_i y_{ci}^2 a_i \alpha b_i}{\sum_i 2\omega_i y_{ci}^2 b_i^2} \tag{22}$$

where

$$\begin{cases} A = \sum_i 4\omega_i b_i^2 x_i^2, D = \sum_i 4\omega_i b_i \\ B = \sum_i 2\omega_i b_i^2, E = \sum_i 4\omega_i \\ C = \sum_i 2\omega_i a_i b_i d_{eye}, F = \sum_i 2\omega_i a_i d_{cam} \end{cases} \tag{23}$$

The optimal value of the scaling factor, S_y, can be computed independently of others, as shown by Eq. (22). However, due to the dependency of Eqs. (19-2 1), the optimal parameter values for d_{cam}, K, and S_x can not be uniquely determined, similar to the case of pointwisc optimization. To solve this problem, we can introduce additional performance criteria such as depth distortion, image resolution, and stereo overlap for optimization. One simple way to solve. the problem is to decide the baseline length, d_{cam}, first based on its physical constraints and its effect on depth distortion and stereo overlap. Then, the scaling factor, S_x, and the shifting parameter, K, can be determined from Eqs. (19)(20) or (19)(21) as follows:

$$S_x = \frac{(C - 2FD) d_{cam}}{A + d_{cam}^2 \left(B - \frac{D^2}{2E}\right)}$$

$$K = \frac{d_{cam}^2 D (C - 2FD)}{2E\left(A + \left(B - \frac{D^2}{2E}\right) d_{cam}^2\right)} - \frac{G}{F} \tag{24}$$

6. Simulation and Analysis

Section 6.1 shows the results of the optimization process in obtaining a good view of a selected area of a simulated manipulator. Section 6.2 analyses the size of the virtual objects as a function of distance from the camera. Section 6.3 considers the cases when the view is most natural for a given magnification.

The simple translation factor 't_z' can also be viewed as a translation component '*dvobj - (a*dobj)*'. The latter definition is more intuitive when tbc viewer is interested in magnifying his view while looking at a specific object located at distance *dobj* from the camera while keeping tbc virtual object at distance *dvobj* from the viewer in the virtual world,

6.1 Simulation

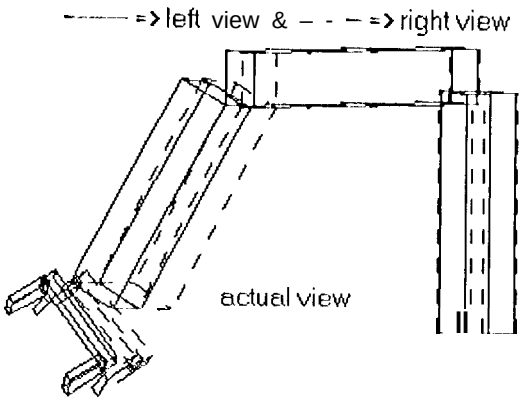

Fig 7. Actual view dobj=1000cm. Optimization for end effector points.

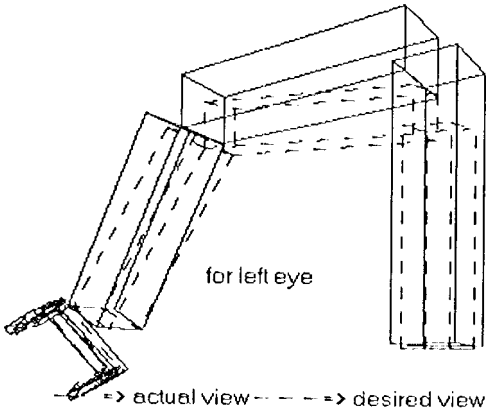

Fig 8. Comparison dobj = 90cm. Optimization for end effector points.

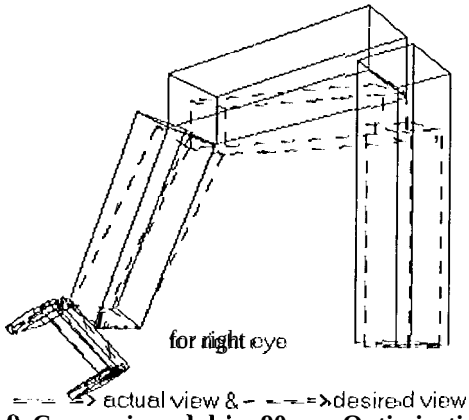

Fig 9. Comparison dobj = 90 cm. Optimization for end effecter points.

because the distance of these points in the desired virtual world becomes negative. What we need to understand from the graph is that if we do not perform the optimization process constantly then as the object of interest moves along z it will appear to shrink or grow. This is clearly uncle.sirable.

In the. second graph on relative magnification dobj = 90 & dvobj = 90. We notice that for a = 0.95 the lnagnification is almost constant and should not requite constant updating. Actually this is noticed because the case is close to the ideal case that will be discussed in the next section. Also if $dvobj > a*(dobj + f)$ *(i.e.* $t_z > a*f$*)* then objects far ther away in the virtual world appear smaller. in the. other case objects farther away will look larger.

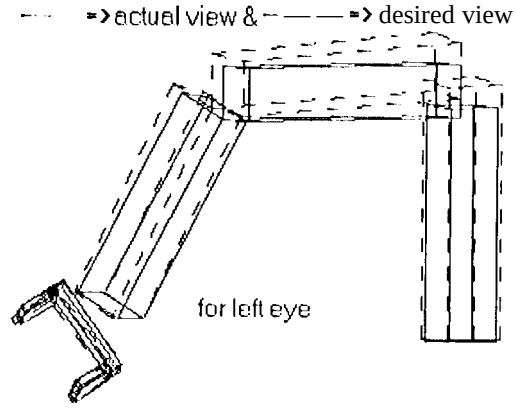

Fig 10 Comparison dobj = 1000 cm. (right eye case is similar & hence not shown).Optimization for end effector points.

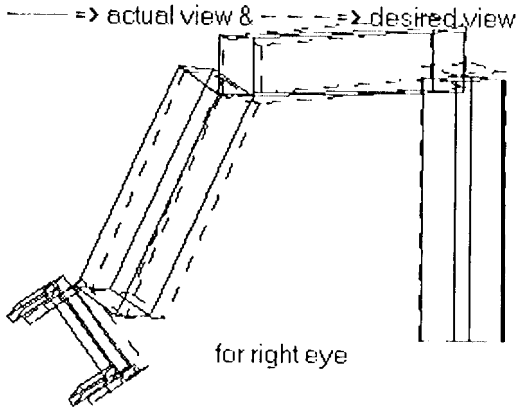

Fig 11. Comparison dobj = 1000 cm (left eye case is similar & hence not shown). Optimization for points on the first link.

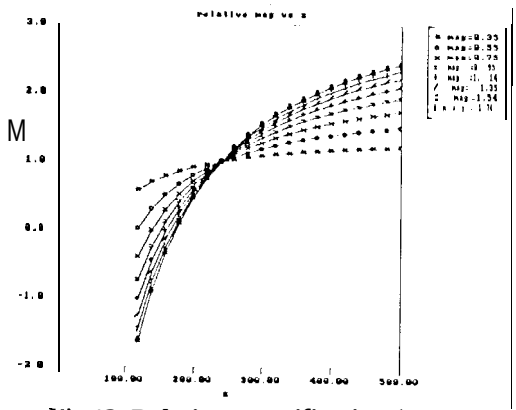

Fig 12. Relative magnification (case 1) z is in crns & α = mag

The usefulness of the optimization process is evident from the results of the simulation. We chose to simulate a manipulator with three links and an end effector. Each of the three links has a single block. The end effector is mode.llcd to look like a gripper. The size of the three links are about 10x50 x10c111[3] each. The size of the end effector is about 20x I 5x10 c11i[3]. The orientation of the manipulator is such that the end effector is farther away from the first link by about 45cm. We perform the optimization process for multiple points on the cad effecter and compare. the results of the actual view obtained with the desired view. Say we decide to place the virtual manipulator at a distance of *dvobj* = 45 cms from the viewer with a magnification of 0.15. Say screen width is 30cm. 'l'his would mean that the screen will hold about 30/0. 15 = 200cm of the real scene's width at the position of the real object,

For the viewer to notice a scaled manipulator at *dvobj*, the view on the screen should be as seen in fig. 4. Fig. 5 shows the actual view obtained by optimizing for the view for pre-selected points on the end effector when the distance of the manipulator from the camera (dobj) is 200cm. Fig. 6 shows the actual view as described for fig.5 for the case when dobj . 90cm. Fig. 7 shows the actual view as described for fig. 5 for the case when dobj = 1000cm. For a better comparison of the results wc show the desired and actual view for various cases separately for each eye. Fig. 8 shows the left eyes view for the optimized case as well as the desired case. We can now clearly see how the actual view matches the desired view better at the end effecter. For this case dobj = 90cm. Fig. 9 shows the views as described for fig. 8 for the right eye. Fig. 10 shows the views as described for fig. 8 for the case when dobj = 1000 cm. We can compare the result with that of optimizing for- all the points on the first link. Fig. 11 compares the right eyes desired view with that of the right eye's actual view. For this case the optimization was done for points on the first link. Wc can see that unlike the other cases the end effector's desired and actual views arc not properly matched but that the first links actual and desired views are properly matched. Hence we can conclude that if wc optimize the viewing parameters for the end effector them it is scaled as specified and placed at tbc specified distance in the perceived world, and if the optimization is done for the first link then the first link is scaled as specified and placed at the specified distance in the perceived world, with negligible errors. The usefulness of the optimization process in obtaining a good view of a selected area or object in the real world is evident from the simulation.

6.2 Relative Magnification

By studying the relative magnification 'M' of the points as a function of z, we can understand the need for continually performing the optimization process We define M as follows:

$$M = \frac{P_y^{SL}}{P_y^{VSL}} \qquad (25)$$

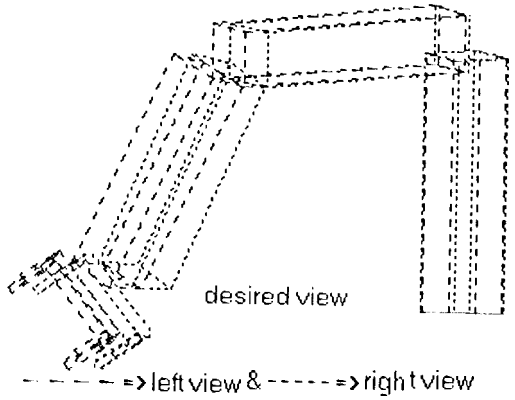

Fig 4. desired view dvobj=45cm, mag= 0.15

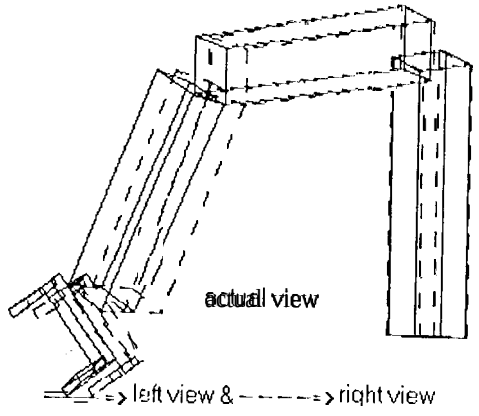

Fig 5. Actual view dobj = 200cm. Optimization for end effector points.

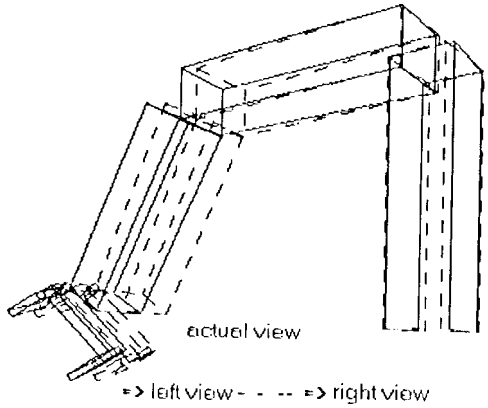

Fig 6. Actual view dobj = 90cm. Optimization for end effector points.

M is independent of y. Consider the graphs in figs. 12 and 13. In the first case, we optimized f for dobj = 200 cm and dvobj = dsc = 45 cm. As noticed from the graph the best value of M (i.e. M = 1) is not at dobj. 'l'his is because it was optimized for a number of points on the simulated manipulator. We also notice that for some value.s of mag M becomes negative. This is

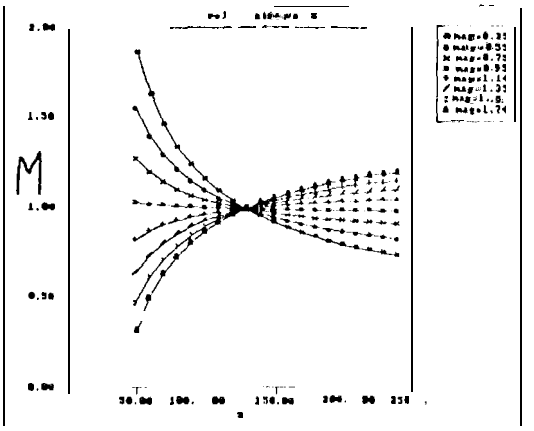

Fig 13. Relative Magnification (case 2)

6.3 **Natural** View

The natural view refers to the case where all of the scene is exactly as it should be if it were magnified by a factor a. We need:

$$M = \frac{P_{y}^{SL}}{P_{Y}^{VSL}} = 1 \qquad (26)$$

If wc have:

$$f + dobj = \frac{dvobj}{\alpha} \qquad (27)$$

then the equation becomes independent of z. i.e. when $t_z = a * f$. But wc still need to have $M = 1$. Assuming no digital scaling, this would necessitate

$$\frac{S_W}{W} = \frac{f}{dsc} \qquad (28)$$

where SW is the screen width and W is the width of the image plane that is mapped onto the screen.

This basically implies that the field of view that is subtended by the screen on the viewer should be equal to the field of view of the camera, Also to have proper shift of the on-screen images and to have uniform variation of the. virtual space. we need to have

$$\begin{cases} d_{cam} = \frac{d_{eye}}{\alpha} \\ k = 0 \end{cases} \qquad (29)$$

Under these conditions the view will be most natural. 1 t is like making the distance between our eyes equal to d_{eye}/a, and then looking at the world. As long as the above parameters arc satisfactory we can stick to the natural view. But in the field of teleoperation versatility anti non-ideal surroundings will require the kind of optimizations presented in this paper.

7. Conclusion

A method of optimizing stereo viewing parameters to present selected areas of an actual scene to look like a user chosen scaled model at a user chosen depth has been presented. convincing simulation results shows the value of the optimization process for real world applications where telepresence is required and/or desired. The optimizing process is not computationally intensive and can hence be used in real time to select new viewing parameters based on required scaling and depth of the virtual world. Coupled with systems that can determine the points on the screen that is seen by the eye., and camera systems that can accurately and rapidly adjust the camera parameters, the optimization process can give the user a feeling of being immersed in the scene with the option of adjusting the scale and relative shift of the virtual world that accurate.ly represents a real world scene.

References

[1] Alberto Rovetta, Remo Sala, Xia Wen, Francesca Cosmi, Arianna Togno and Santo Milanesi "Telerobotic surgery project for laparoscopy" - Robotics 1995 Vol 13, pp397-400.

[2] M. Mallem, F. Chavand and E. Colle "Computer-assisted visual perception in teleoperated robotics - Robotics 1992 Vol 10, pp93-103.

[3] Robert Stone "Virtual Reality and Telepresence" - Robotics 1992 Vol 10, pp461-467

[4] Robert Patterson "Human Stereopsis" - Human Factors 1992, 34(6), pp669-692.

[5] Daniel B. Diner and Derek H. Fender " Human Engineering in Stereoscopic Viewing Devices - Jet Propulsion Laboratory report#JPl,D-8186 15-Jan- 1991.

[6] David F, McAllister "Stereo Computer Graphics and other True 3D Technologies." - Princeton University Press 1993.

Acknowledgement

The research described in this paper was in part carried out by the Jet Propulsion laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.